光变曲线自动抽取程序*

郑捷[1]，江林巧[2]

(1 中国科学院国家天文台，北京 100101；2 乐山师范学院 数理学院，乐山 614000)

摘 要：对各类变星的观测研究是时域天文学中的重要内容，中国科学院国家天文台兴隆观测站作为亚洲大陆规模最大的光学天文观测基地，拥有一批米级口径望远镜，每年面向全球天文工作者开放申请，承担了大量变星的测光观测，但至今没有完全针对兴隆观测站中小口径望远镜的测光数据自动化处理软件。本文介绍了一种针对兴隆观测站望远镜观测数据的光变曲线自动抽取程序，现了从原始观测数据到获得光变曲线的全过程。该程序具备模块化和可自定义等特性，通过简单配置也可适用于其它光学天文望远镜数据。本文还介绍了该程序的结构、特点、使用方法，并以食变星UY UMa的实测数据处理为例展示了处理结果，最后讨论了程序的局限性和对其未来发展的展望。

关键词：光变曲线；自动处理程序；变星；测光数据处理

中图分类号：P111.2　　**文献标识码**：A　　**文章编号**：

1 引言

近年来随着天文观测技术的发展，时域天文观测研究已经成为天体物理学的重要和前沿领域，即通过多波段、时域监测来研究宇宙中各类天体的特征和变化，寻找和探索各类新天体、新现象。在时域天文观测的各类目标中，变星是其中一类重要的观测目标，变星种类众多[1]，对其的观测研究历史悠久，对天文学的发展更是有着强大的推动作用。

中国科学院国家天文台兴隆观测站隶属于中国科学院国家天文台和中国科学院光学天文重点实验室，是亚洲大陆规模最大的光学天文观测基地。兴隆观测站拥有口径大于60厘米的中小口径光学天文望远镜十余台，比如85厘米望远镜、80厘米望远镜、60厘米望远镜等等。这些望远镜每年面向全球天文工作者开放申请，承担了大量变星的观测研究工作。[2]

在对变星的测光观测研究中，对原始数据进行测光处理，获取多波段、高精度的变星光变曲线是一项基础且重要的工作。多年来天文学家们开发了相应的测光数据处理软件包，比如虽然已经停止维护但仍然广泛使用的IRAF[3]软件中自带各种非常成熟的测光软件包。传统的测光数据处理方式需要手动对原始数据进行逐步操作，数量处理效率低下且费事费力。近年来虽然各类光变曲线处理程序不断涌现[4-6]，自动化程度、易用性等较以前有较大幅度提高，但由于不同的望远镜对观测数据文件的命名以及头文件中各类关键词信息等都未作统一，导致很多光变曲线处理程序的通用性十分欠佳，即程序只适用于某一特定望远镜所产生的数据，当将其应用于其它望远镜时，程序常常需要手动更改参数方可继续运行，甚至无法运行。兴隆观测站作为亚洲大陆最大的光学天文观测基地，多台米级口径望远镜每天都在进行大量变星测光观测，但至今还没有完全针对兴隆站米级口径望远镜的测光数据自动化处理软件。

鉴于此，本文针对兴隆米级口径望远镜研发了一套光变曲线自动抽取程序，只要提供原始观测数据的保存目录，该程序即可实现数据类型识别、数据预处理、测光、图像对齐、参考星选取、较差定标等步骤，并自动进行光变曲线绘制。该程序被命名为Quick Light-Curve Pipeline，简称QLCP。该程序已经发布在国内代码托管平台“码云（Gitee）”上。通过该程序的运行，可以让兴隆站的观测者们在对目标进行测光观测的同时，快速对原始数据进行测光处理，获得目标的光变曲线，以便观测者能够准确合理的制定下一步的观测计划，具有一定的科研辅助价值和应用价值。

---

* 基金项目：国家自然科学基金委员会-中国科学院天文联合基金资助（U1631108 ）

收稿日期：2022-07-18；　修订日期：2022-00-00

作者简介：郑捷，男，博士，研究方向：光学实测天体物理，天文软件。 Email: jiezheng@nao.cas.cn

2 程序整体框架

2.1 支撑软件和环境

QLCP程序采用Python3语言进行编写，在找源和测光部分调用Bertin等开发的天文软件Source-Extractor[7]（以下简称SEx）来完成。此外，该程序需要Python天文软件包astropy[8]、pystronomy，以及Python通用软件包：numpy、scipy、matplotlib、os、sys、re等的支持。(由于SEx仅在Linux或macOS操作系统上工作，该程序也受同样限制。)

2.2 程序结构和总体流程

该程序总体结构如图1所示，分为单图像处理和多图像处理两个部分。单图像处理又分为文件类型识别和列表生成、本底合并、平场合并、图像改正、找源和测光这5个模块，而多图像处理则分为图像对齐匹配、参考星选择、较差定标、光变曲线绘制这4个模块。此外程序还包括1个统一调用模块和若干支撑模块。各功能模块之间主要通过输入、输出文件进行耦合。考虑到不同科研课题对数据处理的要求不同，该程序的每个模块均可以使用用户自定义模块替代。

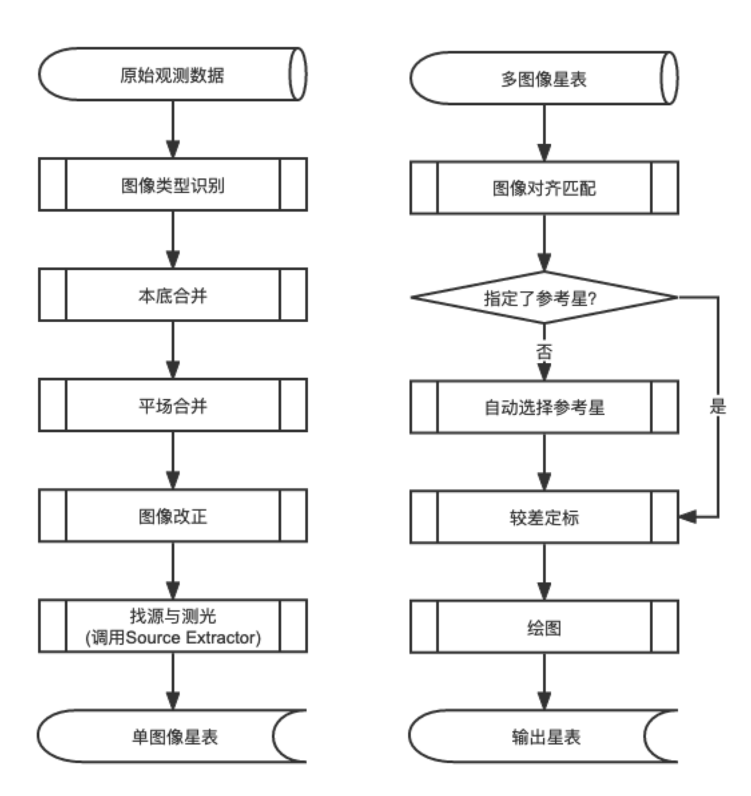


图 1 程序模块示意图，总体可以分为单图像处理（左）和多图像处理（右）两部分

Figure 1 Modules of the program, including the single-epoch image process (left), and the multi-epoch image process (right)

2.3 优势

该程序具备以下几个方面的优势：

（1）自动识别文件：对兴隆站望远镜的观测数据，该程序只需要给定原始数据目录，即可自动识别图像类型，分类进行处理。对于兴隆站之外的其它望远镜数据，也只需要简单配置即可。

（2）图像自动对齐：通过星像匹配算法，高效率、精准实现连续观测图像的对齐工作。

（3）自动选择参考星：如果用户未指定参考星，该程序能够自动从识别到的源中选择参考星，通常为图像中信噪比较高且在观测期间光度稳定的源。

（4）使用方便：只需要指定原始数据保存路径、处理结果保存路径和目标在图像中的位置，

程序即可自动完成光变曲线抽取，也可以进行自定义处理。

（5）模块化与替代性好：该程序各步骤之间通过输出、输入文件进行耦合，任一模块均可以通过自定义操作进行替换，可以局部使用该程序。

（6）自定义性强：该程序可以通过配置文件进行参数自定义，无需修改源代码或编译即可使用自定义参数对数据进行处理，可根据不同的科研需求进行调整，方便应用于不同场合。

3 关键问题和解决

该程序的处理步骤，均参考常用的CCD天文观测数据处理方式[9]，并考虑该程序适用的设备类型做针对性的设计。找源和测光部分调用SEx完成，其它部分均为自行编程实现。其中astropy、numpy、scipy等软件包提供了基础功能支持，例如读写fits文件、数组计算、统计等。以下对部分关键步骤加以说明。

3.1 文件识别

该程序识别观测数据的信息主要依据文件名提供的信息，例如观测类型（本底、平场、科学目标）、观测目标、观测波段、观测序号等，并自动生成文件列表，供后续使用。由于不同望远镜在文件名设置上并不相同，因此该程序通过在配置文件中设置正则表达式进行解析。默认配置文件已经针对兴隆各望远镜进行设置。

如果观测数据文件名中未体现观测信息的数据，也可以通过人工生成文件列表的方式进行处理。人工生成列表还可以实现对数据的筛选，剔除观测错误的图像。

3.2 预处理

识别并生成文件列表后，该程序进行本底和平场合并，并对数据进行改正。本底平场合并均采用中值，而平场在合并之前根据每一幅平场的中值进行归一化。

数据在处理之后，会在指定的输出目录中分类保存。

由于过扫描（overscan）或者预扫描（prescan）区域在许多望远镜探测器中并未包括，因此该程序并未对此进行处理。

此外，目前大部分专业天文望远镜使用CCD或者COMS作为探测器，并通过制冷装置对探测器进行降温，暗流往往处于很低的程度，不需要特意加以改正，因此该程序也并未进行暗流校正。

最后，由于该程序针对多幅测光数据进行处理，宇宙线、运动物体轨迹等干扰因素带来的假源在多幅图像识别中会被自动剔除，该程序也未对此进行改正。

在预处理时，还会进行FITS文件头重建，根据目标信息、观测时间等，计算JD、HJD等信息，以及望远镜观测俯仰角、方位角、月相和月球距离信息等。

3.3 测光

在图像中找源和测光工作，该程序调用成熟的测光软件SEx完成，需要用户自行安装该软件。

SEx需要的配置文件，在该程序中内置了default.sex文件，可以满足大部分情况的需要。如对找源、测光等步骤有特殊配置要求的，可以进行自定义。

SEx输出字段在该程序内置的default.param文件中进行指定，也可以自定义，但至少要求包括目标的中心X、Y坐标、仪器星等和误差等信息，星像的半高全宽（FWHM）和伸长率（Elongation）等信息为可选输出。仪器星等默认选用MAG_AUTO字段，也可以进行自定义。

3.4 图像对齐匹配

处理连续多幅测光图像时，该程序先自动实现图像之间的匹配，从而获取每颗星在不同图像上的位置。该程序在基准图和待匹配图中分别选出最亮的若干颗星，计算两者之间所有星对的x、y坐标距离。由于星像是随机分布的，坐标距离的众值必然是两图之间的实际位移。得到粗略的x、y位移量之后，以此为基础进行进一步目标匹配，得到更加精确的位移量。

基准图像默认为文件列表的第一幅图像，也可以指定列表中任意一幅，甚至可以指定非当晚观测数据作为基准。

3.5 自动选择参考星

对于大部分被观测的变星，观测者往往都会在其周围指定若干颗参考星和检验星。该程序

允许用户手工指定参考星，但是作为一个自动数据处理软件，该程序还可以自动从连续观测的多幅图像中找到亮度稳定的亮星作为参考星。具体做法是，根据图像匹配结果，计算每一幅图像中仪器星等和基准图像的仪器星等的差值，选择其中信噪比较高的星，并剔除异常点之后，计算均值，作为该幅图像和基准图像之间的零点差。随后将该图像仪器星等改正到基准图像的同一水平。最后计算出这些高信噪比星改正后星等在各图之间的标准差，选择标准差最小的，即认为是较稳定的星，作为参考星。

自动选择参考星能够大幅度提升数据处理效率，但是也存在缺陷，参考星的选择只根据当晚的观测数据得出，有可能误选长周期变星。但是这对于生成当天的光变曲线不会有影响。

4 程序使用和性能

4.1 程序发布

该程序已经发布在国内代码托管平台“码云（Gitee）”上，地址为https://gitee.com/drjiezheng/qlcp/。其中包括程序代码、使用说明等。

4.2 使用方式示例

该程序下载后即可放在当前目录下使用。以兴隆站60厘米望远镜在2022年5月14日对食双星UY UMa的V波段观测数据进行处理的为例说明处理过程，原始数据存放于目录/data/raw/20220514_60/，处理结果将保存于/data/raw/20220514_60/目录中，并且通过人工检查第1幅图像，确认目标星中心点大约在(995, 1000)，当晚对该目标进行了8.2小时（0.342天）观测，V波段共取得484幅数据，此外还有10幅本底和5幅晨昏光平场。

在配置文件中，除数据处理选项外，对观测台站和观测设备也进行了定义。根据兴隆站60厘米望远镜的文件命名规则，观测文件命名方式为UYUMa_V_001.fit，用下划线分割目标名称、波段、观测序号等信息，因此相应的配置文件包括以下部分：

```
site_lon = 117.57722     # 117.34.38     #观测站的经度
site_lat = 40.395833     # +40.23.45     #观测站的纬度
site_ele = 960                           #观测站的高度
filename_temp="(?P<obj>[^-_]*)_(?P<band>[a-zA-Z]{0,1})_{0,1}(?P<sn>[0-9]{3}).fit"
```

如果将本程序应用于其它台站，只需要修改相应台站信息即可。根据数据文件名命名规则修改filename_temp关键字指定的正则表达式，例如对于兴隆站85厘米望远镜，其文件名命名方式为：20220514UYUMa-0001V.fit，对应正则表达式为："(?P<date>[0-9]{8})(?P<obj>[^-_]*)-(?P<sn>[0-9]{4})(?P<band>[a-zA-Z]{0,1}).fit"。

以上述60厘米观测数据为例，最简单的调用方式如下。

```
import qlcp22                          # 引入该程序包
qlcp22.do_all(                         # 该程序的统一调用入口
    ini_file="xl60cm.ini",             # 指定目标望远镜配置文件
    raw_dir="/data/raw/20220514_60/",  # 指定原始数据目录
    red_dir="/data/red/20220514_60/",  # 指定输出文件目录
    starxy=[(995, 1000)],              # 指定目标星的位置
)
```

先引入包，随后调用do_all函数自动完成全部步骤。函数参数最少为以上4个，分别是配置文件（ini_file），原始数据文件路径（raw_dir），输出结果路径（red_dir），目标星位置（starxy），其余参数都采用默认配置。将完成从识别文件到最终生成光变曲线的全过程。该程序可选的参数还包括对处理步骤的选择、参考星的指定等，以及对处理步骤的细节进行限制的参数。

整个处理过程约5分钟，最终结果见下一节的介绍。

4.3 输出结果

该程序各个模块之间通过输出文件、输入文件进行耦合，各模块输出文件主要是处理的中间结果。主要输出文件见表1。

表 1 该程序的输出文件

Table 1 Output files of this program

模块	输出文件	内容和格式	格式
文件列表	文件列表	本底列表，各波段平场列表，各波段各目标列表，文件名不包括路径	文本文件
本底平场合并	合并后的本底 合并后的平场	经合并后的本底文件 本底改正并归一化之后合并的各波段平场	简单图像FITS文件 简单图像FITS文件
科学数据改正	改正后科学数据	经本底、平场改正后的科学数据，并计算出观测JD、HJD等信息	简单图像FITS文件
找星和测光	单图测光星表	测光结果，不论采取何种程序测光，统一为该程序内部格式	二进制表格FITS文件
图像对齐	位移量	各图像相对参考图像的x、y方向位移量	文本文件
目标选择	参考星选择结果	如果没有指定参考星，则从数据中选择	文本文件
星表提取	目标星、参考星数据汇总表	根据用户指定或该程序猜测，提取出指定星的星等数据	文本文件和二进制表格FITS文件
较差测光	目标星较差结果	经过较差的目标星光变曲线数据	文本文件和二进制表格FITS文件
图表绘制	光变曲线、证认图	光变曲线和证认图图像	PNG图像文件

4.4 最终输出数据

图2和图3上述数据进行自动处理的结果，其中图2标注了目标星（T-0）和自动选择的参考星（T-1到T-5），图3为得到的目标光变曲线，以及参考星的光变曲线。由于当天观测时气象条件较差，因此得到的光变曲线信噪比不够高。从图中可见，当天的参考星光变标准差在0.005到0.010星等之间。

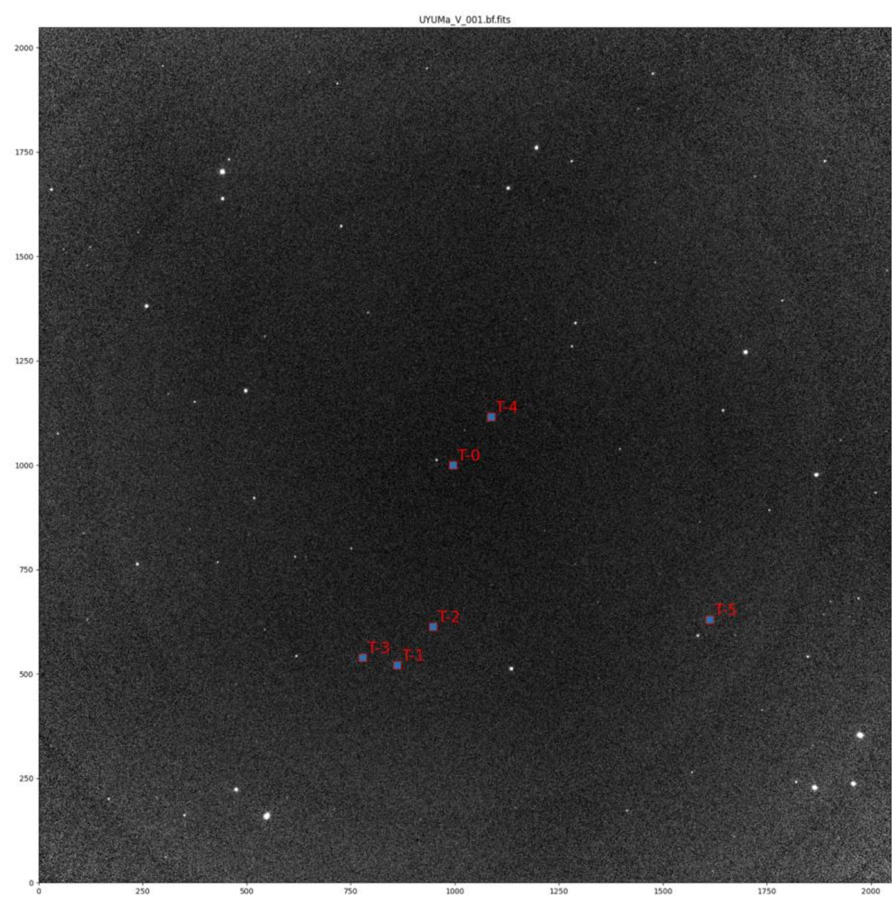


图2 2022年5月14日利用兴隆观测站60厘米望远镜观测食双星UY UMa的V波段图像，图中标注的T-0为目标星，T-1至T-5为自动选择的参考星。

Figure 2 V-band image of the eclipsing binary UY UMa observed by Xinglong 60-cm telescope on May 14, 2022. The mark T-0

indicates the target star, and the marks T-1 to 5 are reference stars chosen by this program.

图3 从2022年5月14日利用兴隆观测站60厘米望远镜观测食双星UY UMa的V波段数据中抽取出的光变曲线，其中红色方块所示曲线为目标星光变曲线，星号所示曲线为参考星光变曲线，编号顺序与图2中所示一致。

Figure 3 V-band light-curve extracted from data of the eclipsing binary UY UMa observed by Xinglong 60-cm telescope on May 14, 2022. The red square marked curve shows the light-curve of the target star, while asterisk curves shows the reference stars, which have the same number in Figure 2.

4.5 存在的问题

该程序只是后续综合程序的基本雏形，针对兴隆观测站米级望远镜的变星观测数据处理进行编制，通过简单配置也可以适用于其它望远镜的变星观测数据处理。在考虑了特定使用场景的情况下，该程序的功能有一定的局限性，将在未来版本中加以改进。

首先，该程序只考虑各幅观测数据之间的平移而不考虑视场的旋转，因此该程序不适用于发生场旋的观测。

此外，该程序并未进行天测定标，因为对于目标明确的变星观测，对星图中目标的具体赤经、赤纬并没有特定的要求，天测等问题将另行解决。

同样的，由于变星研究通常对目标的视星等并没有特别要求，因此该程序也未进行流量定标。而且，实践中较差定标精度实际上往往高于绝对流量定标。

最后，由于该程序使用了被广泛应用的测光软件SEx进行找源和测光，能力和精度受SEx的限制。

总结

本文介绍了从变星连续测光观测数据中抽取光变曲线的程序QLCP，主要包括程序的结构、简单使用方法、处理结果等。该程序可以大幅度简化实测数据的处理，并且具备模块化、可定制等特点。但是该程序作为一个针对特定使用场景程序，有一定的局限性和不足，将在后续的研发中继续改善。

致谢

本文得到国家自然科学基金委员会-中国科学院天文联合基金资助（U1631108）。感谢兴隆60厘米、80厘米、85厘米等望远镜全体工作人员的支持。本文部分工作得到中国科学院光学天文重点实验室开放课题资助。感谢国家天文台王汇娟、王佳琪、罗常青，云南天文台戴智斌等人在试用该程序时提出的意见和建议。

参考文献：

[1] 沈强生. 变星的分类[J]. 云南天文台台刊, 1979, 1:45-52.
Qiangsheng Shen. Classifications of variable stars[J]. Publications of Yunnan Observaroty, 1979, 1:45-52.
[2] 国家天文台兴隆观测站[EB/OL]. http://www.xinglong-naoc.cn/, 2022-07-01.
[3] H. Butcher , R. Stevens. Image Reduction and Analysis Facility Development[J]. Kitt Peak National Observatory Newsletter, 1981, 16:6
[4] 钟文杰, 等. 1.26米红外望远镜测光处理管线设计与实现[J]. 天文研究与技术, 2018, 15(4):465-472.
Zhong Wenjie, et al. Design and Realization of Photometric Processing Pipeline for 1.26m Infrared Telescope[J]. Astronomical Research & Technology, 2018, 15(4):465;472.
[5] David Motl, et al. C-Munipack[EB/OL]. http://c-munipack.sourceforge.net/, 2010
[6] 郑捷, 等. SAGE 测光巡天数据处理方法研究[J]. 天文研究与技术, 2019, 16(1):93-106.
Zheng Jie, et al. Research on the Data Reduction of the SAGE Photometric Survey[J]. Astronomical Research & Technology, 2019, 16(1):93-106.
[7] E. Bertin & S. Arnouts. SExtractor: Software for source extraction[J]. Astronomy & Astrophysics Supplement, 1996, 117:393.
[8] The Astropy Collaboration, et al. The Astropy Project: Building an inclusive, open-science project and status of the v2.0 core package[J]. The Astronomical Journal, 2018, 156(3):123.
[9] M. Bessell. Two-dimensional CCD photometric data reduction[J]. Aust. J. Astron., 1996, 6(5-6):219-222.

An automatic light-curve extraction pipeline

Jie Zheng[1], Lin-qiao Jiang[2]
(1, National Astromonical Observatories, Chinese Academic of Sciences, Beijing 100101, China; 2 School of Mathmetics and Physics, Leshan Normal University, Leshan 614000, China)

Abstract: The observation and study of all kinds of variable stars are important part of time domain astronomy. Xinglong Observatory, NAOC as the largest observatory in the Asia continent. It owns several meter-aperture astronomical telescopes, which are opened to scientists from all countries. There telescopes are observing a lot of variable stars, but a dedicated data reduction pipeline is lacking. This paper introduces an automatic pipeline, which can extract light-curve from raw observation data. This pipeline is highly modularized and customizable, and can be applied to data from different optical astronomical telescopes after minor customization. The structure, the feature, and the usage of this pipeline are discussed in this paper, together with a sample of eclipsing-binary UY UMa. In the end, the limit of this pipeline is listed with future improve planning.

Key words: light-curve, automatic pipeline, variable star, photometric data process